National Défense
Defence nationale




# ESTIMATION OF DOPPLER FREQUENCY IN THE PRESENCE OF NOISE (U)

by

Patrick Lie Chin Cheong



## DEFENCE RESEARCH ESTABLISHMENT OTTAWA

TECHNICAL NOTE 93-15

Canada

September 1993
Ottawa

National Defence / Défense nationale

# ESTIMATION OF DOPPLER FREQUENCY IN THE PRESENCE OF NOISE (U)

by

**Patrick Lie Chin Cheong**
*Electronic Support Measures Section*
*Electronic Warfare Division*

**DEFENCE RESEARCH ESTABLISHMENT OTTAWA**
TECHNICAL NOTE 93-15

PCN
011LB

September 1993
Ottawa

## ABSTRACT

This paper discusses the estimation of the Doppler frequency of an incoming RF signal that is corrupted by additive white Gaussian noise. The estimation is performed using the sampled in-phase and quadrature components of the demodulated signal. The Cramér-Rao lower bounds for the minimum variance estimators as well as the maximum-likelihood (ML) estimator are derived. The mean-square error of the ML estimator is then studied.

## RÉSUMÉ

Cet article discute de l'estimation de la fréquence Doppler d'un signal RF corrompu par du bruit blanc Gaussien et additif. L'estimation s'est faite en utilisant les composantes en phase et en quadrature du signal démodulé. Les bornes inférieures de Cramér-Rao (C-R) pour les estimateurs à variance minimale ainsi que l'estimateur à probabilité maximale sont dérivés. L'erreur quadratique moyenne de l'estimateur à probabilité maximale est ensuite étudiée.

## EXECUTIVE SUMMARY

The problem of estimating the Doppler frequency of a received RF signal is basically the problem of estimating the shift in frequency of the signal in the presence of noise generated in the propagation medium and in the measuring systems. This paper establishes bounds for the best possible frequency estimate that can be achieved from a finite number of I and Q samples. The latter are obtained by demodulating the received signal in its in-phase (I) and quadrature (Q) components which are then discrete-time sampled.

The equations describing the Cramér-Rao (C-R) lower bounds for the minimum variance of the unbiased estimates of the frequency, amplitude, and phase of the signal are derived. It is found that the C-R bound for the frequency estimate is inversely proportional to the signal-to-noise ratio *SNR*, the square of the sampling time-interval $T_s$, and the cube of the number of samples $N$ obtained. The minimum variances for the coefficients of a general polynomial phase signal, i.e, chirp signals, are also derived. The maximum-likelihood (ML) estimation of these parameters are then considered since the ML estimator is the minimum variance estimator. The analysis of the ML frequency estimator, however, reveals that at low SNR, the mean-square error (MSE) of the ML estimate increases much more rapidly than the expected C-R bound due to the higher probability of the occurrence of an outlier. The SNR dependent deviation of the MSE from the C-R bound is the threshold effect. The relationship between the ML estimation and the discrete Fourier transform (DFT) of the sampled complex I/Q components is also shown.

An algorithm for the implementation of the ML frequency estimator is then considered. It consists of two parts; a coarse search and a fine search. The coarse and the fine search are performed using the fast Fourier transform (FFT) and a numerical procedure such as the bisection method or the secant method, respectively. Errors due to the direct estimation of the frequency parameter by the FFT algorithm are also analysed. For moderate to high SNR (above the threshold effect), the dominant error in the frequency estimation is found to be the approximation of the ML estimate by the maximum magnitude DFT point. This approximation can be made more accurate by increasing the number of DFT points, but once this number is chosen, the MSE of the DFT approximation is constant and is independent of the input SNR.

Some numerical examples about the best possible frequency estimates that can

be achieved for coherent and noncoherent pulse radars are finally given. The typical values chosen for the pulse duration, sampling rate, and SNR are $1\,\mu s$, $10^8\,Hz$, and $20\,dB$, respectively. It is seen that in the case of coherent pulses, for the frequency estimate to have an accuracy of $10\,Hz$ or less, the number of pulses required is around 50. On the other hand, in the noncoherent case, for the same accuracy, the required number of pulses is in the order of a 100,000.

# TABLE OF CONTENTS

## LIST OF FIGURES

## LIST OF TABLES

# 1.0 INTRODUCTION

The measurement of the doppler frequency is basically the measurement of the shift in frequency of the received signal. In the case of a radar signal, the doppler frequency shift is related to the radial component of the relative velocity of the target by [9]

$$f_d = -\frac{2\dot{R}f}{c} = -\frac{2\dot{R}}{\lambda} \tag{1}$$

where $\dot{R}$ is the range rate, $\lambda = f/c$ is the transmitted wavelength of the radar signal and $f_d$ is the doppler shift which is positive for decreasing range $R$. For a passive intercept receiver such as the Electronic Intelligence (ELINT) or the Electronic Support Measure (ESM) system, the doppler shift is given by [12]

$$f_d = -\frac{\dot{R}}{\lambda} \tag{2}$$

where $\dot{R}$ is the relative radial velocity between the emitter and the receiver.

There are many reasons why one would use a passive system such as the ESM receiver instead of an active radar to measure the doppler frequency. Among them, one obvious reason is the unavailability of a radar at the time due for example to its use in tracking other targets. More importantly, however, the ESM system has the advantage that it can detect an emitting target much earlier than a radar because of the one way wave propagation instead of two. Thus, information about a change in the target's movement or trajectory can be made available much sooner using an ESM system. Since the intercept (ESM) receiver has no knowledge of the transmitted frequency and can only measure the received signal, it cannot calculate the shift in frequency due to the relative motion between the emitter and the receiver. It can, however, determine the doppler shift due to the change in the radial velocity of the emitter. This can be useful in detecting the radial acceleration or the change in trajectory of an active emitter. As an example, suppose that the velocity of a target emitter (missile) is $V_i$ $m/s$ and is travelling directly towards the receiver (ship A) as shown in Figure 1. At some point along the route it changes trajectory and turns $\phi$ degrees away from the ship. The change in radial velocity due to the change of $\phi$ degrees from the original trajectory is $V_i - V_i \cos\phi = V_i(1 - \cos\phi)$ $m/s$. Hence, the shift in frequency of the signal at the receiver is

$$f_d = -\frac{V_i(1 - \cos\phi)}{\lambda} \tag{3}$$

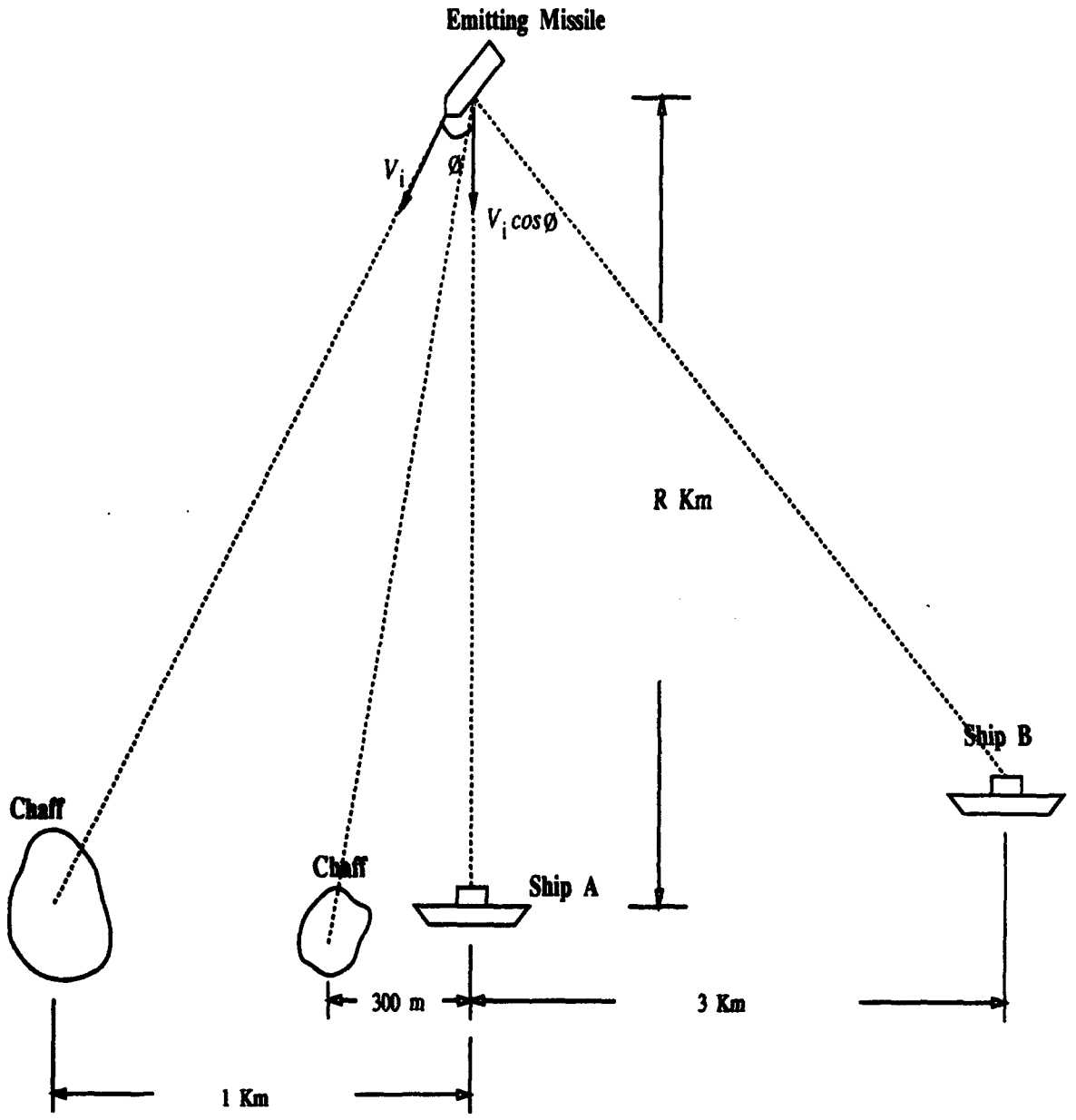


**Figure 1: Change in trajectory of the target emitter.**

For $\phi = 5°$, $\lambda = 0.03\ m$ and $V_i = 300\ m/s$, the doppler frequency shift $f_d$ is then equal to 38.05 *Hz*. In terms of the angle change in trajectory, $\phi$, we can rewrite equation (3) as

$$\phi = \cos^{-1}\left(1 - \frac{f_d \lambda}{V_i}\right) \tag{4}$$

where $f_d$ is the shift in frequency and $f_d \geq 0$. Assuming that we are able to estimate $f_d$ down to say around 10 *Hz*. With the same values of $V_i$ and $\lambda$, it means that we will be able to detect a change in trajectory of at least 2.56°.

Now, we will discuss the advantage of using an ESM system with the capability of measuring the doppler frequency with an accuracy of about 10 *Hz* or less compared to a monopulse radar system that has an angular resolution of 0.3 degrees. Referring to Figure 1, consider the first scenario where the missile has locked onto the ship A and the latter is trying to transfer the lock from itself onto the nearby chaff cloud through the use of some electronic countermeasures. The chaff was released by the ship in an attempt to seduce the missile and as such, it is not very far from the ship. In this scenario, it is assumed that the distance ($d$) between the chaff cloud and ship A is around 300 metres. To simplify the example the ship and chaff cloud are in a line perpenticular to the initial missile trajectory. It is shown that for a receiver which can measure a shift in frequency down to 10 *Hz*, it can detect a change in trajectory of at least 2.56 degrees for $V_i = 300\,m/s$ and $\lambda = 0.03m$. Hence, this implies that the receiver will be able to detect any deviation of the missile onto the chaff within the range $R_{\text{esm}} = d/\tan\phi = 0.3/\tan(2.56°) = 6.703\ Km$.

For a monopulse radar system with an angular resolution of $\beta$ degrees, the maximum range $R$ for which a change in trajectory of $\phi$ degrees can be detected in a time $t_r$ seconds or less is

$$R = \frac{t_r V_i \sin\phi}{\tan\beta} \simeq \frac{t_r V_i \sin\phi}{\beta}, \tag{5}$$

where for small values of $\beta$, $\tan\beta \simeq \beta$. As can be seen from the above expression, the range $R$ is proportional to the required response time $t_r$ and the velocity $V_i$ of the target, but inversely proportional to the angular resolution. A plot of range (R) versus the response time ($t_r$) for different velocities ($V_i$) is shown in Figure 2 with the trajectory change of $\phi = 5°$ and an angular resolution of $\beta = 0.3°$. In the first scenario as described above, the deviation angle $\phi$ must satisfy the equation $\tan\phi = d/(R + t_r V_i \cos\phi)$. Solving simultaneously this equation with that of (5), we obtain the maximum range $R_{\text{radar}} = 5.12\ Km$ for $V_i = 300\,m/s$, $t_r = 2\,s$, $\beta = 0.3°$, and $d = 300\,m$.

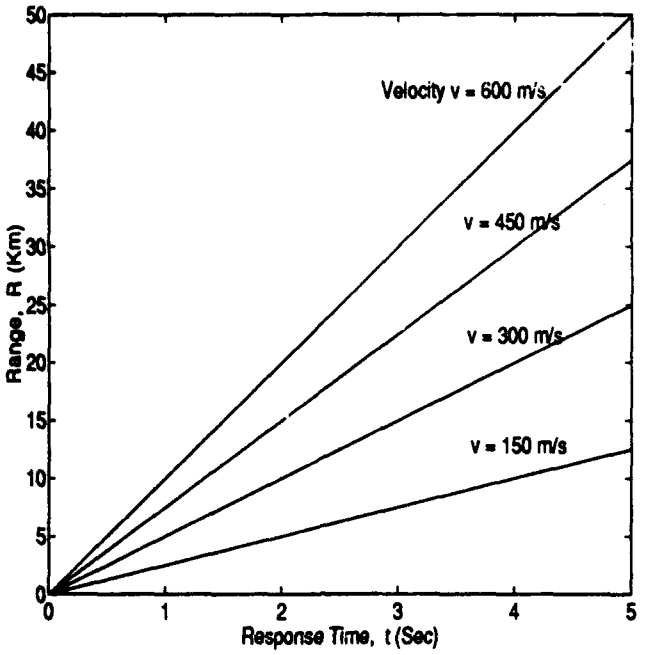


**Figure 2:** Plot of the range $R$ versus the response time $t_r$ for a trajectory deviation of $\phi = 5^o$ and an angular resolution of $0.3^o$.

In the second scenario, the chaff is used against the acquisition mode. That is, the chaff is placed in advance before the missile acquires its target and as such, it is much further away from the ship than in the seduction mode. In this case the distance $d$, between the chaff and the ship is taken to be approximately 1 $Km$ as shown in Figure 1. Using this value of $d$ and with the other variables having the same values as before, the maximum range $R_{\text{esm}}$ for which the ESM doppler receiver can detect a change in trajectory is 22.37 $Km$. Note that $R_{\text{esm}}$ is proportional to the distance $d$ between the ship and the chaff. For the monopulse radar with the same angular resolution of 0.3 degrees and maximum response time of 2 seconds, by solving the previous equations, the maximum range $R_{\text{radar}}$ obtained is 10.39 $Km$.

Similarly, in the third scenario where the missile deviates to track ship B instead of ship A, the doppler receiver can detect this change well beyond the horizon. With the distance $d$ between ship A and B being around 3 $Km$, and with all the other variables remaining the same, we get $R_{\text{esm}} = 67.10$ $Km$ and $R_{\text{radar}} = 18.13$ $Km$. The results are summarized in Table 1 for the three different scenarios.

Ideally, if we were able to determine any shift in frequency at the receiver, we

**Table 1:** Comparison between an ESM Doppler Receiver and a Monopulse Radar System.

| Scenarios | ESM Receiver with Doppler measurement of an accuracy of 10 $Hz$ | Radar System with angular resolution of $0.3^o$ and a required maximum response time of $2\,s$ |
|---|---|---|
| 1. $d = 300\ m$ | $R_{\text{esm}} = 6.710\ Km$ | $R_{\text{radar}} = 5.118\ Km$ |
| 2. $d = 1.0\ Km$ | $R_{\text{esm}} = 22.367\ Km$ | $R_{\text{radar}} = 10.388\ Km$ |
| 3. $d = 3.0\ Km$ | $R_{\text{esm}} = 67.099\ Km$ | $R_{\text{radar}} = 18.130\ Km$ |

would be able to calculate any trajectory change of the target. However, this is usually not the case due to the presence of noise in both the propagation medium and in the systems. This paper addresses the problem of estimating the frequency of the signal which is corrupted by additive white Gaussian noise. The signal is assumed to be of the form $x(t) = a(t)\cos(\omega_d t + \theta)$, where the amplitude $a(t)$, the phase $\theta$, and the angular frequency $\omega_d$ are all assumed to be unknown. In addition, the estimation is performed on a finite number of discrete-time observations.

The minimum variance or mean-square error (MSE) of the unbiased frequency estimation is derived using the Cramér-Rao (C-R) bound. This enables us to have an idea about the best possible accuracy that can be achieved as a function of the number of samples and input signal-to-noise ratio (SNR). To further generalize this concept, we have also included signals that are frequency modulated such as the chirp signals. The lower variance C-R bounds for the estimation of the coefficients of an arbitrary polynomial phase signal of the form $x(t) = a(t)\cos(b_0 + b_1 t + b_2 t^2 + \cdots + b_p t^p)$ are derived. The maximum-likelihood (ML) parameter estimation of the demodulated signal is considered in this paper because it is known that for large values of SNR, the ML estimator is unbiased, and is the minimum-variance estimator whose mean-square error is given by the C-R bound. As the SNR decreases, however, the increase in the probability of the outliers of the frequency estimate occuring, will cause the MSE of the ML estimator to increase much more rapidly than the C-R bound. The tendency of the MSE to start deviating from the C-R bound as the SNR decreases below a certain value is known as the threshold effect. Most nonlinear estimators suffer from threshold effects. It will be also shown that the ML estimation of the frequency parameter is related to the discrete

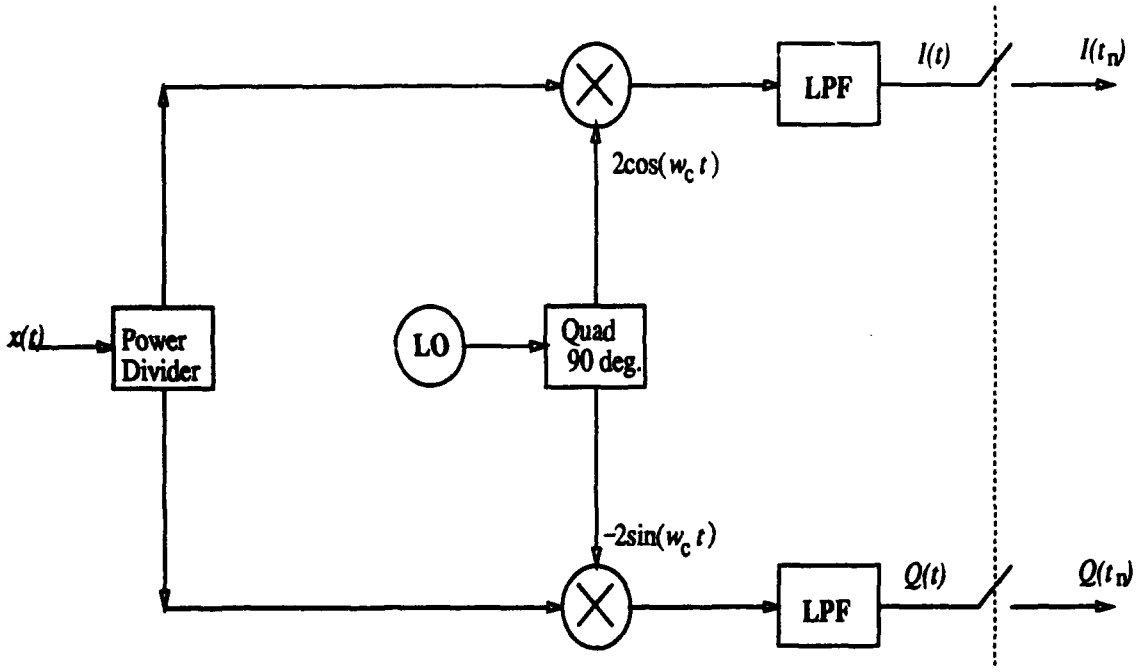


**Figure 3: Block diagram of the in-phase and quadrature (I/Q) demodulator.**

Fourier transform (DFT) of the complex envelope of the sampled signal. The overall MSE in using the DFT as an approximation to the ML frequency estimation is then derived.

## 2.0 COHERENT RECEIVER

We will first assume that the incoming RF signal from the emitter is coherent. It can be a continuous wave (CW), a train of coherent pulses, or any transmitted signal whose carrier is modulated by a waveform $a(t)$. For now, we will further assume that the initial carrier frequency $\omega_c$ has already been obtained and is known. Suppose that there is a shift in frequency of the signal due to the doppler effect. The incoming signal $x(t)$ is then given by

$$x(t) = a(t)\cos[(\omega_c + \omega_d)t + \theta_o] + n(t) \tag{6}$$

where $w_d = 2\pi f_d$ is the doppler shift, and $n(t)$ is the additive white Gaussian noise with zero mean and variance $\sigma^2$. The received signal $x(t)$ is then applied to the in-phase and quadrature (I/Q) demodulator as shown in Figure 3. Note that in practice, the RF signal is usually down-converted to an intermediate frequency (IF) signal before it is applied to the I/Q demodulator. This is not shown here for brevity.

The noise $n(t)$ can be represented as a band noise signal in terms of its in-phase and quadrature components $n_c(t)$ and $n_s(t)$, respectively, where $n_c(t)$ and $n_s(t)$ are also Gaussian with zero mean and variance $\sigma^2$ [4], that is,

$$n(t) = n_c(t)\cos(\omega_c t) - n_s(t)\sin(\omega_c t). \tag{7}$$

The demodulated in-phase and quadrature baseband signals $I(t)$ and $Q(t)$ are then given by

$$I(t) = a(t)\cos(\omega_d t + \theta_o) + n_c(t) \tag{8}$$

$$Q(t) = a(t)\sin(\omega_d t + \theta_o) + n_s(t). \tag{9}$$

The baseband signals $I(t)$ and $Q(t)$ are sampled at a rate of $f_s = \frac{1}{T_s}$ per second. The sampled signals obtained are then equal to

$$I_k = a_k\cos(\omega_d t_k + \theta_o) + n_c \tag{10}$$

$$Q_k = a_k\sin(\omega_d t_k + \theta_o) + n_s, \tag{11}$$

where $t_k = t_0 + kT_s$, for $k = 0, 1, 2, \ldots$ and $I_k$, $Q_k$, $a_k$, $n_c$ and $n_s$ are defined as $I(t_k)$, $Q(t_k)$, $a(t_k)$, $n_c(t_k)$ and $n_s(t_k)$, respectively. The noise samples $n_c$ and $n_s$ are statistically independent Gaussian random variables with mean zero and variance $\sigma^2$. For CW radar signals the amplitude $a(t)$ is usually a constant and for pulsed radar signals the value of $a(t)$ is either a constant or zero, neglecting the rise and fall time of the pulses. In the latter case, we can always sample the signals $I(t)$ and $Q(t)$ such that the value of $a_k$ is the constant but not equal to zero. Thus, we can replace $a_k$ simply by the parameter $a_o$.

## 3.0 VARIANCE BOUNDS ON PARAMETER ESTIMATION

Before the estimations of the parameters such as the frequency $\omega_d$ and the phase $\theta_o$ are made, it is usually very useful to have numbers or values that will indicate the best possible estimates that can be achieved with the available data. Suppose that we have $N$ sample values of $I_k$ and $Q_k$, i.e., $k = 0, 1, \cdots, N-1$. From these values of $I_k$ and $Q_k$, we can form the $N$ sampled complex envelope $Z_k = I_k + jQ_k$, where $j = \sqrt{-1}$. Since the noise samples are independent Gaussian random variables with mean zero and variance $\sigma^2$, the joint probability density function (PDF) of the random variables $Z_k$, $k = 0, 1, \cdots, N-1$, or of the random vector $Z = [Z_0\ Z_1\ Z_2\ \cdots\ Z_{N-1}]$, is the product of the two $N$-variate Gaussian distribution of the random vectors $I = [I_0\ I_1\ I_2\ \cdots\ I_{N-1}]$

and $Q = [Q_0\ Q_1\ Q_2\ \cdots\ Q_{N-1}]$. The PDF of vector $Z$ is, therefore, given by [7]

$$p(Z;\ \omega, a, \theta) = \left(\frac{1}{2\pi\sigma^2}\right)^N \exp\left[-\frac{1}{2\sigma^2}\sum_{k=0}^{N-1}(I_k - \mu_k)^2 + (Q_k - \nu_k)^2\right] \tag{12}$$

where the means $\mu_k$ and $\nu_k$ are as follows:

$$\mu_k = a\cos(\omega t_k + \theta) \tag{13}$$

$$\nu_k = a\sin(\omega t_k + \theta). \tag{14}$$

As can be seen from the above equation (12), the PDF of the random vector $Z$ is a function of the unknown parameters $\omega$, $a$, and $\theta$.

### 3.1 Cramér-Rao Bounds for Frequency and Phase Estimates

Due to the fact that the received signal is corrupted by noise and is, therefore, random in nature, we can only obtain estimations of the signal parameters. Furthermore, since the values of the estimates are dependent on the observation data, they are also random variables and as such have a mean and a variance. An *unbiased* estimator is one whose mean value is equal to the true value of the parameter and a *minimum-variance* estimator is one whose variance is less than or equal to the variance of any other estimator. The minimum-variance unbiased estimator is unique, but it may not always exist. Under certain conditions, it can be shown that the variance of any unbiased estimator cannot be less than a particular lower bound known as the Cramér-Rao (C-R) bound [11, 10].

From [11], the Cramér-Rao bound for an unbiased estimator of a single unknown parameter $\alpha$ is derived and is shown to be

$$\mathrm{Var}(\tilde{\alpha}) = \sigma_{\tilde{\alpha}}^2 \geq \frac{1}{E\left(\left[\frac{\partial \ln p(Z;\alpha)}{\partial \alpha}\right]^2\right)} = \frac{-1}{E\left(\frac{\partial^2 \ln p(Z;\alpha)}{\partial \alpha^2}\right)} \tag{15}$$

where $\tilde{\alpha}$ is the estimate of the true value of $\alpha$. For the variance of the simultaneous estimates of two or more unknown parameters $\alpha_1, \alpha_2, \cdots, \alpha_n$, the unbiased C-R bounds are obtained from the inverse of the $(n \times n)$ Fisher information matrix $\Gamma$ whose elements $\gamma_{ij}$ are given by

$$\gamma_{ij} = E\left\{\frac{\partial \ln p(Z;\ \alpha_1, \alpha_2, \cdots, \alpha_n)}{\partial \alpha_i} \cdot \frac{\partial \ln p(Z;\ \alpha_1, \alpha_2, \cdots, \alpha_n)}{\partial \alpha_j}\right\}$$

$$= -E\left\{\frac{\partial^2 \ln p(Z;\ \alpha_1, \alpha_2, \cdots, \alpha_n)}{\partial\alpha_i \partial\alpha_j}\right\}. \tag{16}$$

If we denote the inverse of the matrix $\Gamma$ by $\Psi$, i.e., $\Psi = \Gamma^{-1}$, then the C-R bound for the variance of the individual unbiased estimate is obtained from the diagonal elements of $\Psi$. That is,

$$\mathrm{Var}(\tilde{\alpha}_i) = \sigma^2_{\tilde{\alpha}_i} \geq \psi_{ii} \tag{17}$$

where $\psi_{ii}$ is the $i$th diagonal element of the matrix $\Psi$.

Now, returning to the case where the PDF of the random variable vector $Z$ is as given in (12) and the unknown parameters $\alpha_1$, $\alpha_2$, and $\alpha_3$ are $\omega$, $a$, and $\theta$, respectively, we have a $(3 \times 3)$ matrix $\Gamma$ whose elements are

$$\gamma_{ij} = \frac{1}{\sigma^2}\sum_{k=0}^{N-1}\left(\frac{\partial\mu_k}{\partial\alpha_i}\right)\left(\frac{\partial\mu_k}{\partial\alpha_j}\right) + \left(\frac{\partial\nu_k}{\partial\alpha_i}\right)\left(\frac{\partial\nu_k}{\partial\alpha_j}\right) \tag{18}$$

where $\mu_k$ and $\nu_k$ are as shown in (13) and (14), respectively. Solving for the previous expression (18), and substituting $t_k = t_0 + kT_s$, the symmetric matrix $\Gamma$ is reduced to

$$\Gamma = \frac{1}{\sigma^2}\begin{bmatrix} a^2\left(Nt_0^2 + N(N-1)t_0T_s + \frac{1}{6}N(N-1)(2N-1)T_s^2\right) & 0 & a^2\left(Nt_0 + \frac{1}{2}N(N-1)T_s\right) \\ 0 & \Lambda & 0 \\ a^2\left(Nt_0 + \frac{1}{2}N(N-1)T_s\right) & 0 & a^2N \end{bmatrix}. \tag{19}$$

Taking the inverse of $\Gamma$, the Cramér-Rao bounds for the unknown parameters $\omega$, $a$ and $\theta$ are as follows:

$$\mathrm{Var}(\tilde{\omega}) = \sigma^2_{\tilde{\omega}} \geq \frac{12\sigma^2}{a^2T_s^2N(N^2-1)} \tag{20}$$

$$\mathrm{Var}(\tilde{a}) = \sigma^2_{\tilde{a}} \geq \frac{\sigma^2}{N} \tag{21}$$

$$\mathrm{Var}(\tilde{\theta}) = \sigma^2_{\tilde{\theta}} \geq \frac{2\sigma^2\left(6t_0^2 + 6(N-1)t_0T_s + (N-1)(2N-1)T_s^2\right)}{a^2T_s^2N(N^2-1)} \tag{22}$$

Note that unlike the frequency $\omega$ and the amplitude $a$, the variance bound on the phase $\theta$ is dependent on the time at which the first sample is taken, i.e., $t_0$. These results were similarly obtained by Rife [8]. By differentiating equation (22) with respect to $t_0$ and setting the result to zero, it is found that the minimum bound is achieved when $t_0 = -\frac{1}{2}(N-1)T_s$ and is equal to $\sigma^2/(a^2N)$. In terms of the signal-to-noise ratio which

is defined as $SNR = a^2/(2\sigma^2)$, the Cramér-Rao bound for the variance of the frequency estimate is $\sigma_\omega^2 \geq 6f_s^2/[(SNR)N(N^2-1)]$, where $f_s = 1/T_s$ is the sampling rate.

### 3.2 Cramér-Rao Bounds for Polynomial Phase Modeled Estimates

We can further generalize the incoming signal $x(t)$ as an arbitrary polynomial chirp signal given by

$$x(t) = a\cos\left(b_0 + (\omega_c + b_1)t + b_2t^2 + \cdots + b_pt^p\right) + n(t) \tag{23}$$

where $a$, $b_0$, $b_1$ ..., $b_p$ are the unknown parameters to be estimated. That is, there is now a modulation on the frequency or phase of the signal. Note that for $p = 1$, the signal is basically the same as that in the previous case where the sinusoidal has a constant frequency equal to $b_1$. For $p = 2$, we have a linear chirp or linear frequency modulation (FM) signal, i.e., its frequency varies linearly as a function of time and the slope is given by $b_2$. Higher order chirp signals can be obtained with $p \geq 3$. As before using the same coherent receiver as that described in section 2, we have

$$I(t) = a\cos(b_0 + b_1t + b_2t^2 + \cdots + b_pt^p) + n_c(t) \tag{24}$$

$$Q(t) = a\sin(b_0 + b_1t + b_2t^2 + \cdots + b_pt^p) + n_s(t). \tag{25}$$

The above in-phase and quadrature signals are sampled into the discrete samples $I_k$, and $Q_k$, $k = 0, 1, \ldots, N-1$, respectively. Defining the complex samples $Z_k = I_k + jQ_k$, the joint probability density function of vector $Z$, given the parameter vector $\vec{\alpha} = [a\ b_0\ b_1\ \ldots\ b_p]$, i.e., $p(Z;\ \vec{\alpha})$, is the same as that expressed in (12), except that the means $\mu_k$ and $\nu_k$ are now given by

$$\mu_k = a\cos(b_0 + b_1t_k + b_2t_k^2 + \cdots + b_pt_k^p) \tag{26}$$

$$\nu_k = a\sin(b_0 + b_1t_k + b_2t_k^2 + \cdots + b_pt_k^p) \tag{27}$$

It is to be noted that $\mu_k$ and $\nu_k$ are now functions of the $(p+2)$ unknown parameters $a$, $b_0$, $b_1$, ..., $b_p$. Similar to the previous example, the Cramér-Rao bounds for the various parameters can be obtained from the inverse of the $((p+2)\times(p+2))$ Fisher information matrix $\Gamma$ whose elements $\gamma_{ij}$ are given by equation (18), and where the unknown parameters $\alpha_i$, $i = 1, 2, \ldots, p+2$, are $a$, $b_0$, $b_1$, ..., $b_p$, respectively. Solving for $\gamma_{ij}$,

$i, j = 1, 2, \cdots, p+2$, we have

$$\Gamma = \frac{a^2}{\sigma^2} \begin{pmatrix} \frac{N}{a^2} & 0 & 0 & \cdots & 0 \\ 0 & S_0 & S_1 & \cdots & S_p \\ 0 & S_1 & S_2 & \cdots & S_{p+1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & S_p & S_{p+1} & \cdots & S_{2p} \end{pmatrix} \tag{28}$$

where the elements $S_i$ are

$$S_i = \sum_{k=0}^{N-1} (t_k)^i = \sum_{k=0}^{N-1} (t_0 + kT_s)^i . \tag{29}$$

The matrix $\Gamma$ can now be inverted and the diagonal elements of $\Gamma^{-1}$ yield the C-R lower bounds for the variance of the unbiased estimates of each of the parameters $a$, $b_0$, $b_1$, $\cdots$, $b_p$. Note that the inverse of the matrix $\Gamma$ can be computed by reducing it to the $(p+1)\times(p+1)$ matrix $\Gamma'$, where

$$\Gamma' = \frac{a^2}{\sigma^2} \begin{pmatrix} S_0 & S_1 & S_2 & \cdots & S_p \\ S_1 & S_2 & S_3 & \cdots & S_{p+1} \\ S_2 & S_3 & S_4 & \cdots & S_{p+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ S_p & S_{p+1} & S_{p+2} & \cdots & S_{2p} \end{pmatrix} . \tag{30}$$

Once the inverse of $\Gamma'$ is obtained, then $\Gamma^{-1}$ is simply the augmented matrix of $(\Gamma')^{-1}$ as shown below.

$$\Gamma^{-1} = \begin{pmatrix} \frac{\sigma^2}{N} & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & \Big[ & (\Gamma')^{-1} & \Big] \\ 0 & & & \end{pmatrix} . \tag{31}$$

What is required now is to calculate $(\Gamma')^{-1}$. The matrix $\Gamma'$ has a Hankel type structure and hence, can be efficiently computed using such recursive procedure as the Levinson-Durbin algorithm [3]. We will now consider a $(4 \times 4)$ matrix $\Gamma'$ which implies a third order polynomial phase modeled of the signal, i.e., $x(t) = a\cos(b_0 + b_1 t + b_2 t^2 + b_3 t^3)$. The inverse matrix

$$\Psi' = (\Gamma')^{-1} = \frac{\sigma^2}{a^2} \frac{1}{D} \begin{pmatrix} \varphi_{11} & \varphi_{12} & \varphi_{13} & \varphi_{14} \\ \varphi_{21} & \varphi_{22} & \varphi_{23} & \varphi_{24} \\ \varphi_{31} & \varphi_{32} & \varphi_{33} & \varphi_{34} \\ \varphi_{41} & \varphi_{42} & \varphi_{43} & \varphi_{44} \end{pmatrix} \tag{32}$$

where the determinant $D$ is given by

$$\begin{aligned} D = \; & S_3^4 - 3S_2S_3^2S_4 + S_2^2S_4^2 + 2S_1S_3S_4^2 - S_0S_4^3 + 2S_2^2S_3S_5 - 2S_1S_3^2S_5 - 2S_1S_2S_4S_5 \\ & + 2S_0S_3S_4S_5 + S_1^2S_5^2 - S_0S_2S_5^2 - S_2^3S_6 + 2S_1S_2S_3S_6 - S_0S_3^2S_6 - S_1^2S_4S_6 \\ & + S_0S_2S_4S_6 \end{aligned} \tag{33}$$

and the elements $\varphi_{i,j}$, $i, j = 1, 2, 3, 4$, are

$$\begin{aligned} \varphi_{11} &= S_2S_4S_6 - S_4^3 + 2S_3S_4S_5 - S_2S_5^2 - S_3^2S_6 \\ \varphi_{12} = \varphi_{21} &= S_3S_4^2 - S_3^2S_5 - S_2S_4S_5 + S_1S_5^2 + S_2S_3S_6 - S_1S_4S_6 \\ \varphi_{13} = \varphi_{31} &= S_2S_4^2 - S_3^2S_4 + S_2S_3S_5 - S_1S_4S_5 - S_2^2S_6 + S_1S_3S_6 \\ \varphi_{14} = \varphi_{41} &= S_3^3 - 2S_2S_3S_4 + S_1S_4^2 + S_2^2S_5 - S_1S_3S_5 \\ \varphi_{22} &= S_0S_4S_6 - S_3^2S_4 + 2S_2S_3S_5 - S_0S_5^2 - S_2^2S_6 \\ \varphi_{23} = \varphi_{32} &= S_3^3 - S_2S_3S_4 - S_1S_3S_5 + S_0S_4S_5 + S_1S_2S_6 - S_0S_3S_6 \\ \varphi_{24} = \varphi_{42} &= S_2^2S_4 - S_2S_3^2 + S_1S_3S_4 - S_0S_4^2 - S_1S_2S_5 + S_0S_3S_5 \\ \varphi_{33} &= S_0S_2S_6 - S_2S_3^2 + 2S_1S_3S_4 - S_0S_4^2 - S_1^2S_6 \\ \varphi_{34} = \varphi_{43} &= S_2^2S_3 - S_1S_3^2 - S_1S_2S_4 + S_0S_3S_4 + S_1^2S_5 - S_0S_2S_5 \\ \varphi_{44} &= S_0S_2S_4 - S_2^3 + 2S_1S_2S_3 - S_0S_3^2 - S_1^2S_4 \end{aligned} \tag{34}$$

For simplicity it is assumed that the first sample is taken at $t_0 = 0$, which means that $t_k = kT_s$. In this case, the elements $S_i$ of the Fisher matrix $\Gamma$ becomes $S_i = \sum_{k=0}^{N-1}(kT_s)^i = T_s^i \sum_{k=0}^{N-1} k^i$. For the $(4 \times 4)$ matrix $\Gamma'$, the elements $S_0$ to $S_6$ are as follows:(refer to [1])

$$\begin{aligned} S_0 &= N \\ S_1 &= \frac{N(N-1)}{2} T_s \\ S_2 &= \frac{N(N-1)(2N-1)}{6} T_s^2 \\ S_3 &= \frac{N^2(N-1)^2}{4} T_s^3 \\ S_4 &= \frac{N(N-1)(2N-1)(3N^2-3N-1)}{30} T_s^4 \\ S_5 &= \frac{N^2(N-1)^2(2N^2-2N-1)}{12} T_s^5 \\ S_6 &= \frac{N(N-1)(2N-1)(3N^4-6N^3+3N+1)}{42} T_s^6 \end{aligned} \tag{35}$$

Note that we are only interested in the diagonal elements of $(\Gamma')^{-1}$. The Cramér-Rao lower bounds on the parameter estimates $\tilde{a}$, $\tilde{b}_0$, $\tilde{b}_1$, $\tilde{b}_2$, $\tilde{b}_3$ (the superscript $\tilde{}$ denotes the

estimator of the given parameter) are then given by

$$\mathrm{Var}(\tilde{a}) = \sigma_{\tilde{a}}^2 \geq \frac{\sigma^2}{N} \tag{36}$$

$$\mathrm{Var}(\tilde{b}_0) = \sigma_{\tilde{b}_0}^2 \geq \frac{\sigma^2}{a^2}\frac{\varphi_{11}}{D} = \frac{4(2N-1)(N^2-N+3)}{(SNR)N(N+1)(N+2)(N+3)} \tag{37}$$

$$\mathrm{Var}(\tilde{b}_1) = \sigma_{\tilde{b}_1}^2 \geq \frac{\sigma^2}{a^2}\frac{\varphi_{22}}{D} = \frac{100(6N^4-27N^3+42N^2-30N+11)}{(SNR)\,N(N^2-1)(N^2-4)(N^2-9)\,T_s^2} \tag{38}$$

$$\mathrm{Var}(\tilde{b}_2) = \sigma_{\tilde{b}_2}^2 \geq \frac{\sigma^2}{a^2}\frac{\varphi_{33}}{D} = \frac{180(9N-13)(2N-1)}{(SNR)\,N(N^2-1)(N^2-4)(N^2-9)\,T_s^4} \tag{39}$$

$$\mathrm{Var}(\tilde{b}_3) = \sigma_{\tilde{b}_3}^2 \geq \frac{\sigma^2}{a^2}\frac{\varphi_{44}}{D} = \frac{1400}{(SNR)\,N(N^2-1)(N^2-4)(N^2-9)\,T_s^6} \tag{40}$$

where we have defined the signal-to-noise ratio $SNR = a^2/2\sigma^2$. These expressions can be simplified if we assume that the number of samples $N$ is large, i.e., $N \gg 1$. In this case, the value of $S_i$ reduces to

$$S_i = T_s^i \sum_{k=0}^{N-1} k^i \approx T_s^i \frac{N^{i+1}}{i+1}. \tag{41}$$

Substituting the above equation in (33) and (34), we have the approximate values of the variance lower bounds of the parameters' estimates $\tilde{b}_0$, to $\tilde{b}_3$ as

$$\mathrm{Var}(\tilde{b}_0) = \sigma_{\tilde{b}_0}^2 \geq \frac{8}{(SNR)\,N} \tag{42}$$

$$\mathrm{Var}(\tilde{b}_1) = \sigma_{\tilde{b}_1}^2 \geq \frac{600}{(SNR)\,N(NT_s)^2} \tag{43}$$

$$\mathrm{Var}(\tilde{b}_2) = \sigma_{\tilde{b}_2}^2 \geq \frac{3240}{(SNR)\,N(NT_s)^4} \tag{44}$$

$$\mathrm{Var}(\tilde{b}_3) = \sigma_{\tilde{b}_3}^2 \geq \frac{1400}{(SNR)\,N(NT_s)^6} \tag{45}$$

From the expression of the lower variance bounds for the estimation of the polynomial phase coefficients, it is seen that all the variances are inversely proportional to the signal-to-noise ratio as expected. In addition, it can be observed that the variance bound of the polynomial coefficient is inversely proportional to twice its polynomial power of the time duration $\tau = NT_s$ of the signal, where $N$ is the number of samples and $T_s$ is the sampling

interval. For example, the estimation of the coefficient $b_2$, the second order polynomial, has a variance which is inversely proportional to $\tau^4 = (NT_s)^4$. This means that the higher the order of the polynomial coefficient, the more number of samples or a longer time duration of the signal is required to achieve a certain accuracy in the estimation of that coefficient. It is to be noted that the exact expressions in (37)–(39) depend on the initial sampling time $t_0$. Their minimum bounds are usually achieved when $t_0 = -\frac{(N-1)}{2} T_s$.

## 4.0 MAXIMUM-LIKELIHOOD ESTIMATION

Having determined the lower bounds for the mean-square errors of the parameters' estimates, we will now try to obtain these estimates using the maximum-likelihood (ML) method. The ML estimator maximizes the probability density function $p(Z;\ \vec{\alpha})$ with respect to the vector $\vec{\alpha}$, where the elements of $\vec{\alpha}$ are the parameters to be estimated and $Z$ is the observed sample data. It is to be noted that the maximum of $p(Z;\ \vec{\alpha})$ occurs at the same point as the maximum of $\ln p(Z;\ \vec{\alpha})$. Hence, the maximization of the PDF $p(Z;\ \vec{\alpha})$ is equivalent to the maximization of $\ln p(Z;\ \vec{\alpha})$. We will first consider the case of the simple sinusoid waveform. From (12), we can, therefore, maximize the expression

$$\ln p(Z;\ \omega, a, \theta) = C - \frac{1}{2\sigma^2} \sum_{k=0}^{N-1} \left[ (I_k - \mu_k)^2 + (Q_k - \nu_k)^2 \right] \tag{46}$$

where $C$ is a constant, independent of the parameters $\omega$, $a$, and $\theta$. By noting that the summations of $I_k^2$ and $Q_k^2$ over $k$ are constants once an observation has been made, and eliminating these and some other constants, we obtain a simplified expression to be maximized as

$$L = 2 \sum_{k=0}^{N-1} (I_k \mu_k + Q_k \nu_k) - N a^2 \tag{47}$$

where $\mu_k$ and $\nu_k$ are as defined in (13) and (14), respectively, and $\sum_{k=0}^{N-1} (\mu_k^2 + \nu_k^2) = \sum_{k=0}^{N-1} a^2 = Na^2$. To find the maximum of $L$ we first observe that we can rewrite $(I_k \mu_k + Q_k \nu_k)$ in terms of the real part of the complex envelope $Z_k a e^{-j(\omega t_k + \theta)}$. Substituting $t_k$ by $t_0 + kT_s$ and if normalizing the result by the number of points, N, we get

$$L_o = 2a\, \Re\{A(\omega)\, e^{-j\theta}\, e^{-j\omega t_0}\} - a^2 \tag{48}$$

$$\text{with} \qquad A(\omega) = \frac{1}{N} \sum_{k=0}^{N-1} Z_k\, e^{-jk\omega T_s} \tag{49}$$

where $\Re[\cdot]$ denotes the real part of $[\cdot]$.

Since the parameters $\omega$, $a > 0$, and $\theta$ are all unknown, it can be seen that $L_o$ is maximized over $\theta$ if $\Re\{A(\omega)\, e^{-j\theta}\, e^{-j\omega t_0}\} = |A(\omega)|$. This occurs when $\theta = \arg\{A(\omega)e^{-j\omega t_0}\}$, where arg$[\cdot]$ is defined as the argument or phase of $[\cdot]$. That is,

$$\max_{\theta} L_o = 2a|A(\omega)| - a^2. \tag{50}$$

If $\omega_{\max}$ is the value of $\omega$ which maximizes $|A(\omega)|$, then

$$\max_{\theta,\omega} L_o = 2a|A(\omega_{\max})| - a^2. \tag{51}$$

In order to find the value of $a$ which maximizes $L_o$, the above equation can be differentiated with respect to $a$ and the result set to zero. Performing this, we obtain $a = |A(\omega_{\max})|$, and

$$\max_{\theta,\omega,a} L_o = |A(\omega_{\max})|^2. \tag{52}$$

Hence, the maximum-likelihood estimates for $w_d$, $a_o$, and $\theta_o$ denoted by $\tilde{\omega}_d$, $\tilde{a}_o$, and $\tilde{\theta}_o$, respectively, are

$$\tilde{\omega}_d = \max_{\omega} |A(\omega)| \tag{53}$$

$$\tilde{a}_o = |A(\tilde{\omega}_d)| \tag{54}$$

$$\tilde{\theta}_o = \arg\left\{A(\tilde{\omega}_d)\, e^{-j\tilde{\omega}_d t_0}\right\}. \tag{55}$$

It is to be noted that the optimization of the function $L_o$ can be performed as described above, because $L_o$ is separable in terms of its variables $\omega$, $a$, and $\theta$.

## 4.1 ML Estimation of the Polynomial Coefficients

For the general case of a sinusoid having a polynomial phase model as described in (23), the maximization of $\ln p(Z;\ \vec{\alpha})$, where the parameter vector $\vec{\alpha} = [a\ b_0\ b_1\ \cdots\ b_p]$, can be similarly reduced to the expression $L$ in (47), but with $\mu_k$ and $\nu_k$ given by (26) and (27), respectively. In this case, $(I_k\mu_k + Q_k\nu_k)$ can be rewritten in terms of the real part of the complex envelope $Z_k\, a \exp\{-j(b_0 + b_1 t_k + b_2 t_k^2 + \cdots + b_p t_k^p)\}$. As before, if we assume $t_0 = 0$, then $t_k = kT_s$, and after normalization by $N$ (the number of points), the

expression to be maximized is

$$L_o = 2a\,\Re\{e^{-jb_0}\,B(b_1,\,b_2,\,\cdots,\,b_p)\}\;-\;a^2 \tag{56}$$

where the function

$$B(b_1,\,b_2,\,\cdots,\,b_p) = \frac{1}{N}\sum_{k=0}^{N-1} Z_k\,\exp\left(-j[b_1(kT_s) + b_2(kT_s)^2 + \cdots + b_p(kT_s)^p]\right). \tag{57}$$

Similar to the previous case, $L_o$ is maximized over $b_0$ when $b_0 = \arg\{B(b_1,\,b_2,\,\cdots,\,b_p)\}$. With this value of $b_0$, we have

$$\max_{b_0} L_o = 2a|B(b_1,\,b_2,\,\cdots,\,b_p)| - a^2. \tag{58}$$

If $\hat{b}_1,\,\hat{b}_2,\,\cdots,\,\hat{b}_p$ are the values of $b_1,\,b_2,\,\cdots,\,b_p$ that maximizes $|B(b_1,\,b_2,\,\cdots,\,b_p)|$ and which, therefore, also maximizes $L_o$, then

$$\max_{b_0,b_1,\cdots,b_p} L_o = 2a|B(\hat{b}_1,\,\hat{b}_2,\,\cdots,\,\hat{b}_p)| - a^2. \tag{59}$$

The maximization of the above expression (59) over the parameter $a$ occurs when $a = |B(\hat{b}_1,\,\hat{b}_2,\,\cdots,\,\hat{b}_p)|$ and hence, the maximum of $L_o$ over all the parameters $(a, b_0, b_1, \cdots, b_p)$ is

$$\max_{a,b_0,b_1\cdots,b_p} L_o = |B(\hat{b}_1,\,\hat{b}_2,\,\cdots,\,\hat{b}_p)|^2. \tag{60}$$

As can be seen, this result is a simple generalization of the result in (48)–(52). To solve for the maximum likelihood estimation of the polynomial phase signal, requires the maximization of the function $|B(b_1,\,b_2,\,\cdots,\,b_p)|$ over a $p$-dimensional space, which is a quite difficult problem for large values of $p$. From here on, only the simple case where the signal is a sinusoid having a fixed frequency, i.e., $p = 1$, will be considered.

## 4.2 Relationship with the Discrete Fourier Transform

From the definition of the discrete Fourier transform of the sampled complex envelope $\{Z_k\}$ [5], and normalized by the number of points $N$, we have

$$A_l = \frac{1}{N}\sum_{k=0}^{N-1} Z_k\,e^{-j\frac{2\pi kl}{N}} \qquad \text{for } l = 0,\,1,\,2,\,\cdots,\,N-1. \tag{61}$$

Thus, from the expression of $A(\omega)$ in (49), it can be seen that

$$A_l = A\left(\frac{2\pi l}{NT_s}\right) = A\left(\frac{\omega_s l}{N}\right) \qquad \text{for } l = 0, 1, 2, \cdots, N-1 \tag{62}$$

where $\omega_s = 2\pi/T_s = 2\pi f_s$. The values of $A_l$, therefore, represent the sampling points of the continuous spectrum $A(\omega)$, with the sampling intervals equal to $\omega_s/N$. Note that $A(\omega)$ repeats itself at every $\omega_s$ interval, that is, $A(\omega)$ is periodic with period $\omega_s$. Hence, to find the maximum of $|A(\omega)|$, we need to search only in the interval between 0 and $\omega_s$. In addition, if we have an a priori knowledge about the range or the maximum value of the frequency to be estimated ($\omega_d$), we can reduce the search space by choosing a sampling rate $\omega_s$ close to the largest possible value of $\omega_d$.

Figures 4 and 5 show the spectra of $|A(f)|$ ($f = \frac{\omega}{2\pi}$), with $N = 64$, $f_d = 500\,Hz$, $f_s = 10\,KHz$, and $SNR = 0\,dB$ in Fig. 5 and $N = 32$, $f_d = 500\,Hz$, $f_s = 10\,KHz$, and $SNR = 20\,dB$ in Fig. 6. The crosses along the curves of the figures represent the magnitudes of the N points DFT of $A(f)$, that is, the $\{|A_l|\}$ points, $l = 0, 1, \cdots N-1$. The value of $f$ at which the maximum point of $|A(f)|$ occurs is the ML estimation of $f_d$, i.e., $\tilde{f}_d$, and the value of $|A(f)|$ at that point is the ML estimation of $a_o$, i.e., $\tilde{a}_o$. In the case of Figure 4, the values of $\tilde{f}_d$ and $\tilde{a}_o$ obtained are $\tilde{f}_d = 507.23\,Hz$ and $\tilde{a}_o = 0.961$, whereas those of Figure 5, $\tilde{f}_d = 499.27\,Hz$ and $\tilde{a}_o = 0.985$. From the $N$ DFT points in Figure 4, the maximum is $|A_4|$ which has a value of 0.86 and which is at $468.75\,Hz$ and in Figure 5, the maximum is $|A_3|$ with a value of 0.75 and at $625\,Hz$.

Comparing the two plots, it can be observed that having a larger value of $N$ causes the width of the peaks (lobes) in the spectrum to become smaller, that is, the width is inversely proportional to $NT_s$. As the signal-to-noise ratio increases, the sidelobes' peaks decrease and the values of $\tilde{f}_d$ and $\tilde{a}_o$ approach the exact values as expected. It is to be noted that if no noise is present, $|A(\omega)|$ will be a sinc function which is symmetric about the doppler frequency $\omega_d$ and which has a period of $\omega_s$. The global maximum point will then occur at exactly $\omega_d$ and will have an exact amplitude of $a_o$. In the presence of noise, however, the sinc function of $|A(\omega)|$ is distorted with the sidelobes' peak values becoming larger and the global maximum point being shifted away from the true value of $\omega_d$. The difference between the maximum DFT point to the true maximum of $|A(\omega)|$ depends on $M$, the number of DFT points used in the computation, but is independent of the SNR. The larger the value of $M$, the more accurate is the DFT approximation.

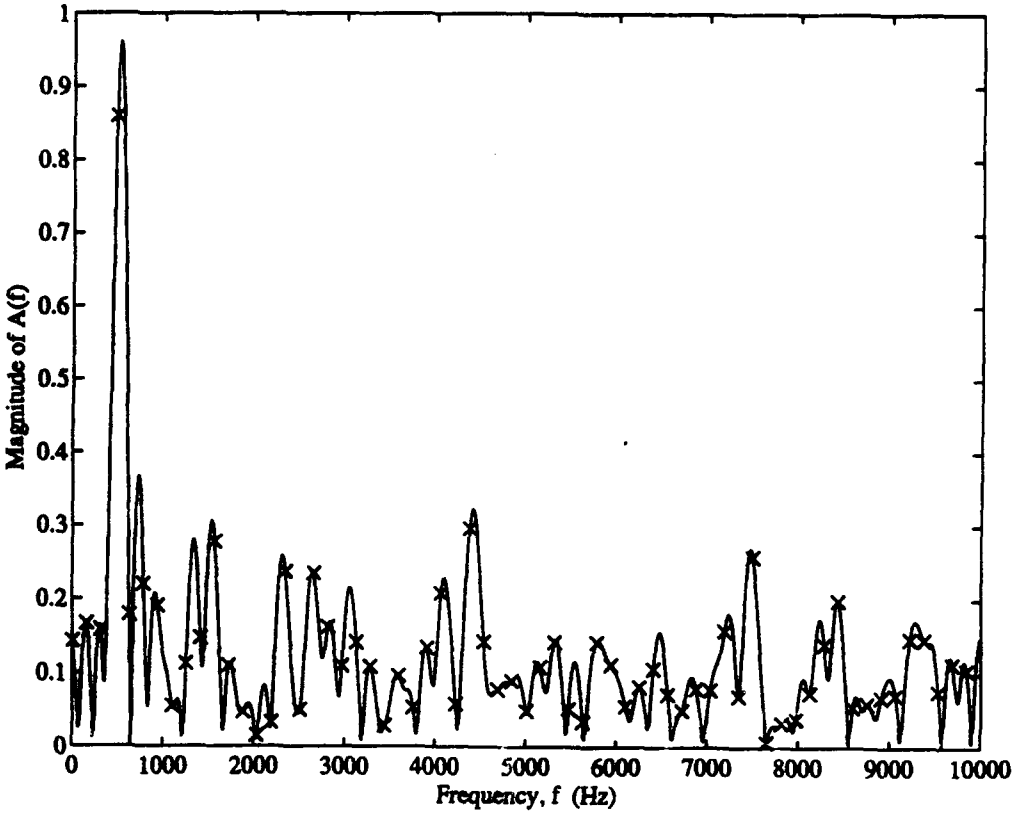


**Figure 4:** Spectrum of $|A(f)|$ with $N = 64$, $f_d = 500\,Hz$, $f_s = 10\,KHz$, and $SNR = 0\,dB$.

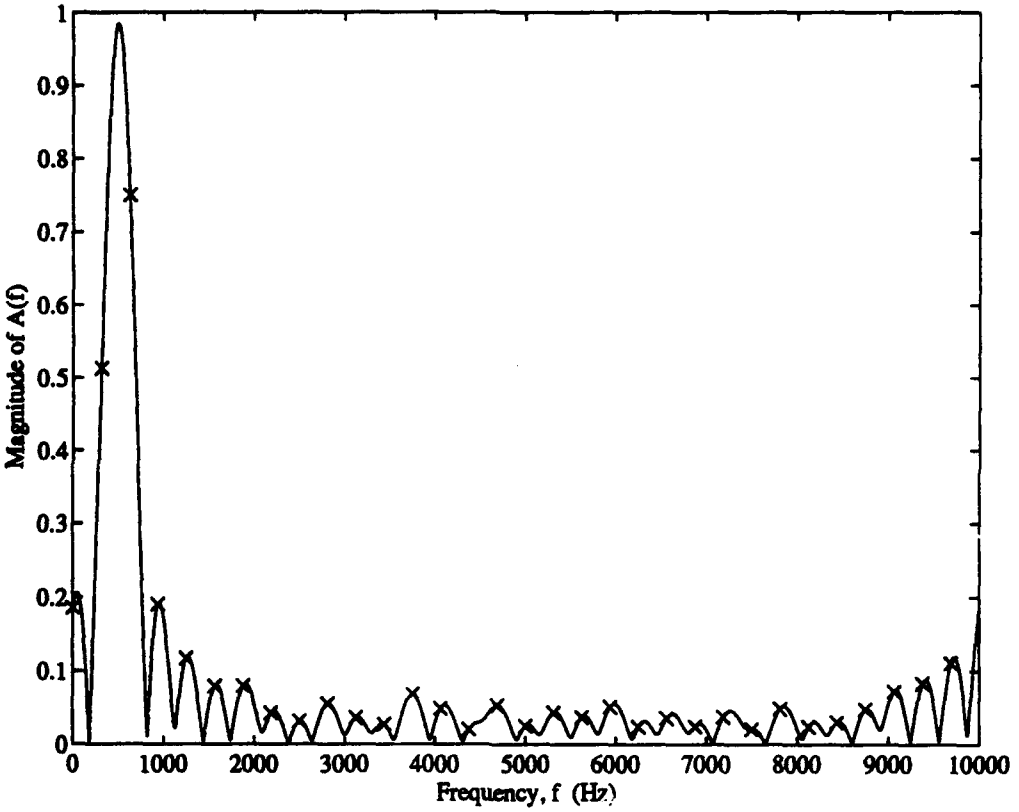


**Figure 5:** Spectrum of $|A(f)|$ with $N = 32$, $f_d = 500\,Hz$, $f_s = 10\,KHz$, and $SNR = 20\,dB$.

### 4.3 Threshold Effects Due to the Occurrence of Outliers

In this section we will consider only the case of the ML estimation of the frequency parameter. As stated previously, for large SNR, the mean-square error of the ML estimation can be obtained from the Cramér-Rao bound given in (20). However, as the SNR decreases, the probability of occurrence of an outlying frequency component whose magnitude is the largest, increases. An outlier is defined as that frequency component which is not in the main lobe containing the transmitted signal frequency but which has a magnitude larger than the main lobe. When the maximum peak frequency component (outlier) occurs far removed from the exact frequency $\omega_d$, severe errors occur in the ML frequency estimate. It is the nature of these large errors caused by the presence of outliers that will be examined below.

For simplicity of the analysis, it will be assumed that an outlier can occur anywhere between zero and $\omega_s$. That is, its probability density function is approximately uniform. In this case, the variance or the MSE due to an outlier is $\omega_s^2/12$. The MSE of the ML estimate for any value of SNR is, therefore, given by

$$\begin{aligned} MSE &= \text{P(outlier)} \cdot \text{(MSE due to outlier)} + \text{P(no outlier)} \cdot \text{(MSE due to no outlier)} \\ &\approx q \cdot \frac{\omega_s^2}{12} + (1-q) \cdot \frac{3\omega_s^2}{2\pi^2 (SNR) N(N^2-1)} \end{aligned} \tag{63}$$

where $q$ is the probability of an outlier and in the absence of any outlier, the MSE is taken from the C-R bound. The expression for $q$ was derived in [8] by defining the random variable $X_l = |A_l|$, for $l = 0, 1, \cdots, N-1$. $X_l$ has a Rayleigh distribution when it is any frequency component other than the frequency of the signal, i.e., the outlier, and has a Rician distribution when it is the signal frequency component. The expression of $q$, thus obtained, was shown to be

$$\begin{aligned} q &= 1 - \int_0^\infty \left[1 - \exp\left(-\frac{Nx^2}{2\sigma^2}\right)\right]^{N-1} \frac{Nx}{\sigma^2} \exp\left[-\frac{N(x^2+a^2)}{2\sigma^2}\right] I_0\left(\frac{Nax}{\sigma^2}\right) dx \\ &= \frac{1}{N} \sum_{i=2}^{N} \frac{N!(-1)^i}{(N-i)!\, i!} \exp\left[-N(SNR)\frac{i-1}{i}\right] \end{aligned} \tag{64}$$

where $I_0(\cdot)$ is the modified Bessel function of the first kind and the signal-to-noise ratio $(SNR) = a^2/(2\sigma^2)$ is as defined before. For large values of $N$, the summation representation in (64) is not easily computed because of the factorial term which can get extremely

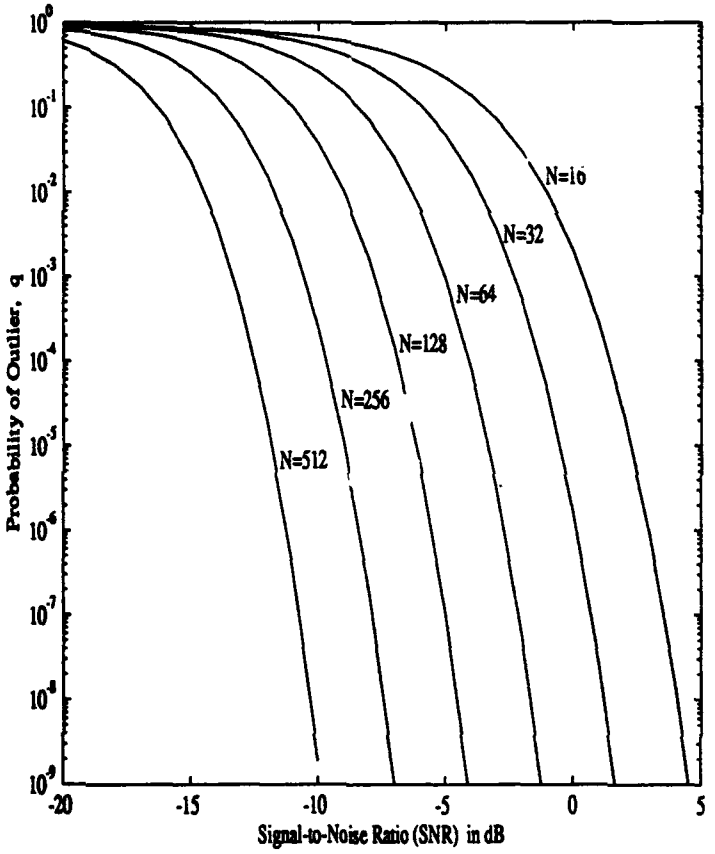


**Figure 6: Probability of outlier, $q$, versus $SNR$ for various values of $N$.**

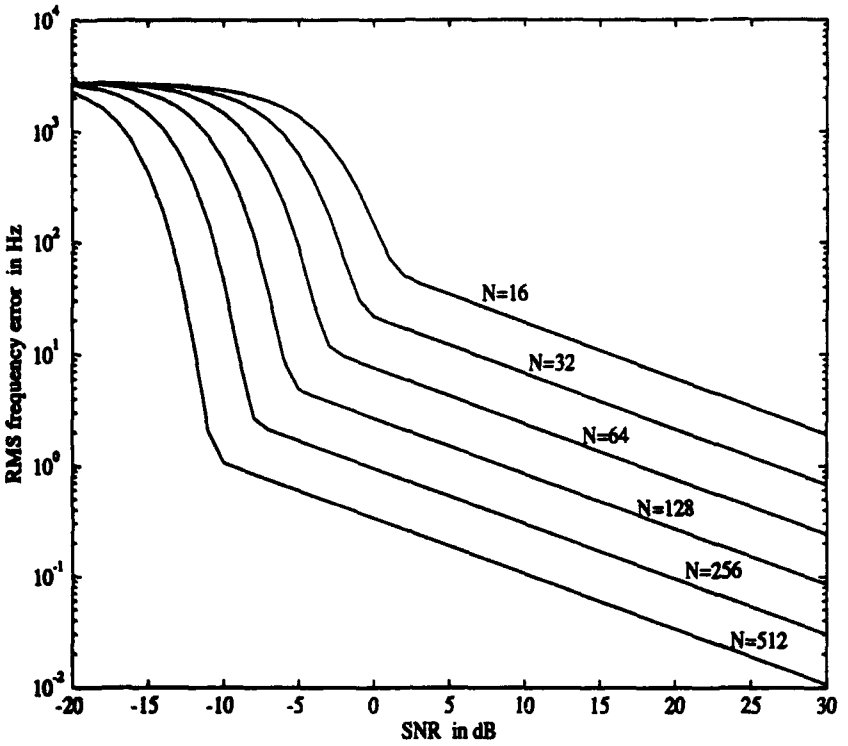


**Figure 7:** Plot of RMS error of the ML frequency estimates against SNR, with $f_s = 10\,KHz$.

large. In these cases, the numerical evaluation of the integral representation is more practical. The plots of $q$ versus $SNR$ for different values of $N$ are shown in Figure 6. With these values of $q$, the mean-square error of the maximum-likelihood frequency estimation given by (63) can be calculated. The root mean square (RMS) error of the ML estimation is then obtained by taking the square-root of the MSE. The RMS frequency errors are plotted against the input SNR for different values of $N$ in Figures 7 and 8, where the sampling rate $f_s = \omega_s/(2\pi)$ are $10\,KHz$ and $100\,KHz$, respectively. As can be seen from the plots of Figures 7 and 8, depending on the values of $N$ and $f_s$, the threshold effect takes place somewhere below $0\,dB$. The straight lines in the graphs represent approximately the C-R bounds. For a given value of $N$, $SNR$, and $f_s$, it can be easily determined if the ML estimation is being operated above or below the threshold effect directly from the plots in Figures 7 and 8. For example, if $N = 128$, $SNR = 20\,dB$, and $f_s = 10\,KHz$, then from the graph of Figure 7, the ML estimation is above the threshold and the RMS error of the frequency estimate is approximately equal to the C-R bound which is 1.69 $rad/s = 0.26\ Hz$.

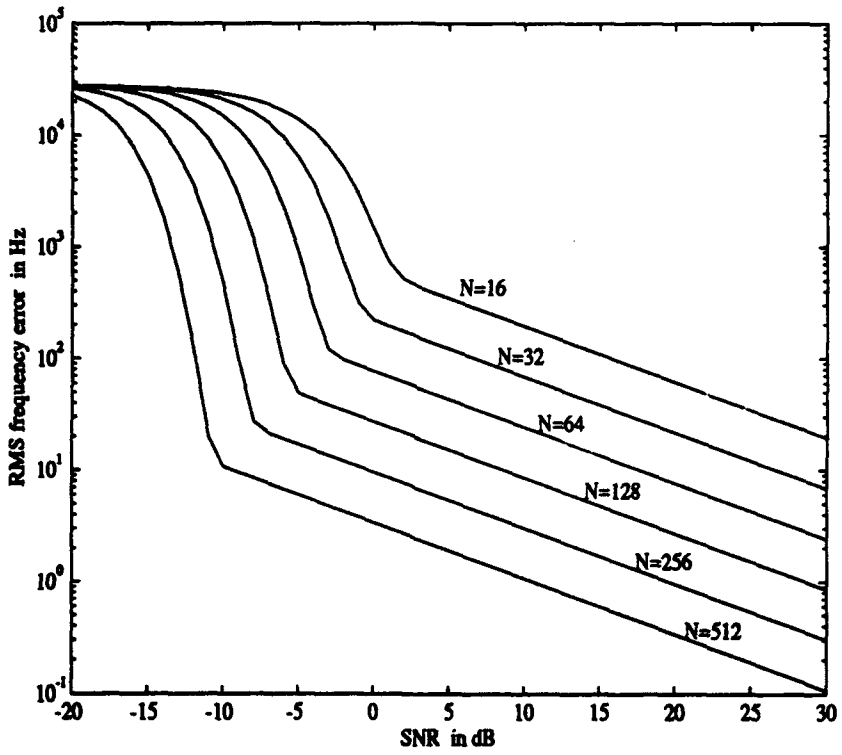


**Figure 8:** Plot of RMS error of the ML frequency estimation against SNR, with $f_s = 100\,KHz$.

## 5.0 DETERMINATION OF THE ML FREQUENCY ESTIMATE

It is shown in Section 4, that the ML frequency estimation of the signal is the angular frequency $\omega$ that maximizes the magnitude of $A(\omega)$, where $A(\omega)$ is as given in (49). To find $\max_{\omega} |A(\omega)|$, we could differentiate the function $|A(\omega)|^2 = A(\omega)A^*(\omega)$ with respect to $\omega$ and set the result to zero. This, however, does not usually work, since $|A(\omega)|$ or $|A(\omega)|^2$ is not a monotonic function of $\omega$ but has numerous local maxima and minima. In order to determine the global maximum of $|A(\omega)|$, we first require a rough or coarse estimate of where it is located. Once this is obtained, a fine search can then be performed to find the exact location of the global maximum point.

The coarse estimate of $\omega_{\max}$, where $\omega_{\max} = \max_{\omega} |A(\omega)|$, can be made directly from the computation of the DFT of the complex data $\{Z_k\}$, $k = 0, 1, \cdots N-1$ as was done in [6]. With the value of $l\omega_s/M$ for which $|A_l|$ is maximum, as an initial guess, the ***bisection*** or the ***secant*** method can be used to determine $\omega_{\max}$ at a desired accuracy. To do this, we need to evaluate the derivative function

$$\begin{aligned} F(\omega) &= \frac{d}{d\omega}[A(\omega)A^*(\omega)] \\ &= \left(A^*(\omega)\frac{d}{d\omega}[A(\omega)]\right) + \left(A^*(\omega)\frac{d}{d\omega}[A(\omega)]\right)^* = 2\Re\left\{A^*(\omega)\frac{d}{d\omega}[A(\omega)]\right\} \end{aligned} \quad (65)$$

where $A^*(\omega)$ is the complex conjugate of $A(\omega)$ and $\Re\{\cdot\}$ is the real part of $\{\cdot\}$. Both the bisection and the secant methods use an iterative procedure to compute $\omega_{\max}$ and at the maximum point, $F(\omega_{\max}) = 0$. The initial interval in which the numerical algorithm is applied is usually taken to be the distance between two consecutive DFT points, i.e., $\omega_s/M$, where $M \geq N$. This will ensure that there exists at most one peak frequency component in this interval, since the widths of the peaks are inversely proportional to $NT_s$.

For both numerical procedures, we need to evaluate $F(\omega_{n+1})$ at each iteration. In the case of the bisection method, $\omega_{n+1} = \frac{1}{2}(c_n + d_n)$, where $c_n$ and $d_n$ are the end-points of the interval after $n$ iterations and for the secant method, $\omega_{n+1}$ is calculated from

$$\omega_{n+1} = \omega_n - F(\omega_n)\frac{\omega_n - \omega_{n-1}}{F(\omega_n) - F(\omega_{n-1})}. \quad (66)$$

The secant method usually has a faster rate of convergence than the bisection method. If

we assume that after $\ell$ number of iterations the numerical procedure is stopped, then the computational complexity is of the order of $3\ell N$ complex multiplications and $2\ell N$ complex additions due to the evaluation of $F(\omega_n)$. To determine the initial coarse estimation, we compute an $M$-point DFT, where $M \geq N$. This can be performed using the fast Fourier transform (FFT) which requires $M \log_2 M$ complex multiplications and additions when $M$ is a power of 2. For moderate to high SNR, the coarse search can usually be done with $M = N$. The overall computational complexity using this method of coarse and fine search is, therefore, of the order of $N(\log_2 N + \ell)$ complex multiplications and additions.

An alternative to the previous search method is to directly find an approximate value of $\omega_{\max}$ by having a very fine sampling of the continuous spectrum of $|A(\omega)|$. This can be achieved by computing a large number of DFT points. That is, an $M$-point FFT is performed on the complex data $\{Z_k\}$, $k = 0, 1, \cdots, N-1$, where $M$ is a multiple of $N$, i.e., $M = mN$, where $m > 1$. In this case the computational complexity is of the order of $M \log_2 M = mN(\log_2 N + \log_2 m)$ complex multiplications and additions. Note that for $M > N$, the $N$ points data $\{Z_k\}$ is padded with trailing zeros to length $M$ before the DFT or FFT is performed. The approximate value of $\omega_{\max}$ is then given by $l_{\max} \cdot \omega_s/M$, where $l_{\max} = \max_l |A_l|$ and $l = 0, 1, \cdots, M-1$.

The exact value of $\omega_{\max}$ can be assumed to be situated anywhere between $l_{\max}$ and its neighboring DFT point. This implies that $\omega_{\max}$ is a random variable that is uniformly distributed in the interval of $\omega_s/M$. Hence, the MSE associated with the approximation of $\omega_{\max}$ by $l_{\max} \cdot \omega_s/M$ is equal to $\omega_s^2/(12M^2)$. The overall mean-square error $\varepsilon^2$ in determining the signal or doppler frequency $\omega_d$ is, therefore, equal to the variance of the ML estimation of $\omega_d$ in the presence of noise as given by (63) plus the variance of approximating $\omega_{\max}$ by the maximum magnitude DFT point. That is,

$$\varepsilon^2 = q\frac{\omega_s^2}{12} + (1-q)\frac{3\omega_s^2}{2\pi^2(SNR)N(N^2-1)} + \frac{\omega_s^2}{12M^2} \tag{67}$$

with $M = mN$, and $m > 1$. For moderate to high signal-to-noise ratio, the variance of the ML estimation can be approximated by the C-R bound and the overall MSE $\varepsilon^2$ can be determined from

$$\varepsilon^2 \approx \frac{3\omega_s^2}{2\pi^2(SNR)N(N^2-1)} + \frac{\omega_s^2}{12m^2N^2}. \tag{68}$$

The overall root-mean-square error $\varepsilon$ is then obtained by taking the square-root of $\varepsilon^2$. Depending on the values of $m$ and $N$ used, it is seen from the above expression (68) that the MSE due to the use of the maximum magnitude DFT point as an approximation to

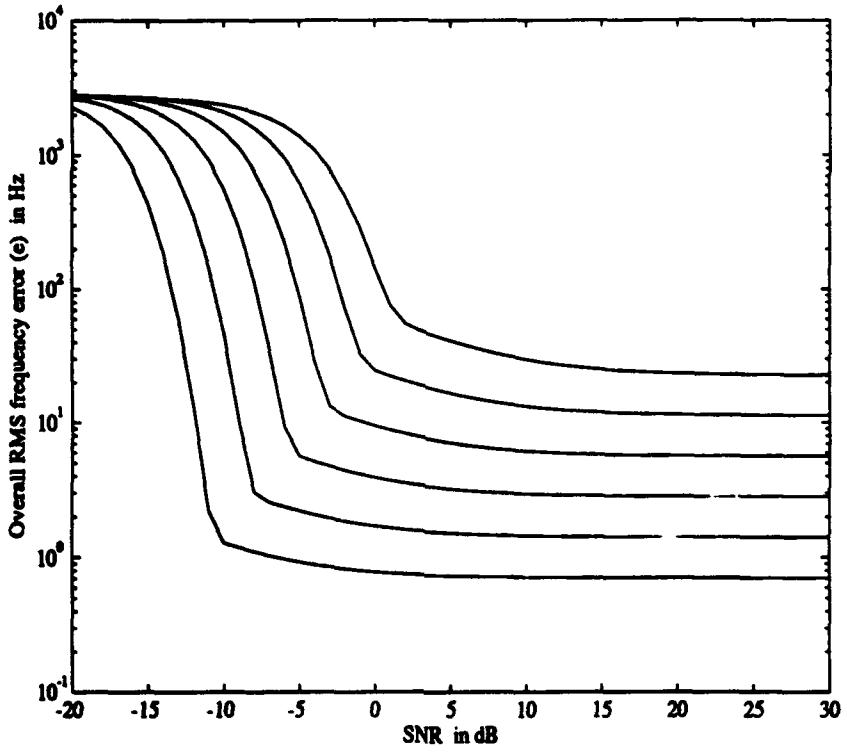


**Figure 9:** Overall RMS frequency error as a function of $SNR$ and $N$ with $f_s = 10\,KHz$ and $m = 8$.

the signal's frequency (second term) can be much larger than the inherent MSE of the ML frequency estimation (first term). At a low SNR, however, the dominant error component is due to the occurrence of outliers given by the first term of Equation (67). The overall RMS error $\epsilon$ is shown in Figure 9 as a function of SNR with the sampling rate $f_s = 10\,KHz$ and the value of $m$ equals to 8. Comparing Figures 7 and 9, the effect due to the MSE of the DFT approximation can be easily observed. As stated earlier, this error increasingly becomes the dominant factor as the SNR increases. This is because the variance of the ML estimation given by the C-R bound is inversely proportional to the SNR, whereas the DFT approximation error is independent of SNR. Hence, with increasing values of SNR, the MSE of the ML estimation becomes negligible compared to the MSE of the DFT approximation which remains basically constant for a given value of $m$ and $N$.

## 6.0 SUMMARY AND CONCLUSION

In this paper, we have studied the problem of estimating the signal parameters, more specifically the doppler frequency, from a finite number of noisy I (in-phase) and Q (quadrature) samples. The received signal is assumed to be coherent and therefore, samples can be taken at regular intervals. In addition, the signal is also assumed to be corrupted by additive white Gaussian noise. The equations that describe the Cramér-Rao bounds which give the minimum variance of an unbiased estimator are derived. In the case of the frequency estimation, the C-R lower bound is shown to be inversely proportional to SNR, the square of the sampling time-interval $T_s$ and the cube of the number of samples $N$ observed. As for the phase estimation, it is found that the C-R bound is dependent on the time $t_0$ at which the first sample is taken. The lower variance bound is also derived for a more generalized chirp signal whose phase is modeled as a polynomial of order $p$, where $p > 1$. The example of the C-R bounds for the coefficients' estimates of a cubic polynomial ($p = 3$) was described in detail.

The maximum-likelihood estimation for the unknown parameters of both the simple sinusoid and the chirp signal are derived. The relationship of the ML frequency estimate to the discrete Fourier transform is then shown. The ML estimator is the minimum variance estimator but, however, its analysis reveals that at low SNR, it is also plagued by the threshold effect. That is, there is a range of low SNR for which the MSE starts increasing very rapidly as the SNR decreases. This is because of the large error committed by the ML estimator in the presence of an outlier. The threshold behaviour of the ML estimation was discussed in Section 4.3.

The algorithm for the implementation of the ML frequency estimator was then considered. It consists of two parts; a coarse search and a fine search. The coarse search yields a rough estimate of where the global maximum of the spectrum is located and this is performed using the fast Fourier transform algorithm. The fine search uses the coarse estimate as the initial guess in order to give a more accurate result of the frequency estimation. It can be performed using numerical procedures such as the bisection and the secant methods. Errors due to the direct estimation of the frequency parameter by the FFT algorithm was also analysed. It was found that at moderate to high SNR (above the threshold effect), the dominant factor in the overall MSE of the frequency estimation is the DFT approximation to the ML estimate. The accuracy of this approximation can be increased by increasing the number of DFT points but at the expense of more number of computations. Once the number of DFT points is chosen, the MSE due to the DFT

approximation is constant and independent of the input SNR.

We have determined the minimum variance frequency estimation using the ML estimator. From the expression of its C-R bound, it is seen that this estimation can be improved by increasing the effective time duration, $NT_s$, of the received signal. The longer the time duration, the more accurate the frequency of the signal can be estimated. Using the proposed method for estimating the frequency of the signal during a single pulse of, say, one microsecond, with a sampling rate of $10^8\,Hz$, and a $SNR = 20\,dB$, the best RMS error that can be achieved is $3.899\,KHz$. Since the doppler frequency expected in practice is of the order of a few hundred hertz or less, it is obvious that this method cannot measure or even detect any doppler shift from a single received pulse. The large value of the estimation error is due to the small pulse time duration. It is to be noted that for a given time duration $\tau$, where $\tau = NT_s$, the accuracy of the minimum variance estimation can be increased by increasing $N$, the number of samples, for the same value of $\tau$. This is because, from the expression of the C-R bound in (20), the MSE is inversely proportional to the cube of $N$ ($N^3$), whereas it is only inversely proportional to the square of the sampling interval $T_s$ ($T_s^2$). This reduction in MSE for the ML estimation, however, will be apparent only for high SNR (above the threshold effect) and it will be of almost no effect on the overall MSE if the DFT approximation is used.

## 6.1 Pulsed Signals

Finally, some examples will be considered, where the target emitter is transmitting either a burst of coherent or noncoherent pulses at the receiver. The goal is to determine the best possible estimation that can be achieved with a finite number of pulses. It is assumed that the received pulses have a moderate to high SNR so that the maximum likelihood estimator is operating above the threshold value.

In the case of the coherent pulses where basically a continuous wave signal is being modulated by a square waveform, suppose that $P$ number of pulses, each with a time duration of $\tau$ seconds, are received. The pulsed signal is sampled at a rate of $f_s = 1/T_s\,Hz$ which implies that the number of samples per pulse is $N = \tau/T_s = \tau f_s$. The sampling of the coherent signal can be performed continuously from pulse to pulse such that the integration over the $P$ pulses is equivalent to the sampling of a CW signal of time-length $P\tau$ seconds. The MSE of the ML angular frequency estimation is then

obtained from

$$\frac{6}{(SNR)T_s^2 NP(N^2P^2-1)} \approx \frac{6}{(SNR)T_s^2 N^3 P^3} \qquad \text{for } N,\ P \gg 1. \tag{69}$$

Comparing the above equation with that of (20), it is seen that by having $P$ number of coherent pulses, the MSE is improved by a factor of $P^3$. Thus, in terms of the RMS error, the improvement factor is of the order of $P^{3/2}$. For the example used earlier, where $\tau = 1\,\mu s$, $f_s = 10^8\,Hz$, and $SNR = 20\,dB$, the increase in accuracy of the frequency estimation when $P = 10$ is by a factor of 31.62 that yields an RMS error of $123.28\,Hz$. Similarly, if we require the frequency estimate to have an RMS error of less than $10\,Hz$, then for the same values of $\tau$, $f_s$, and $SNR$, the number of coherent pulses $P$ needed to achieve this accuracy is $P \geq 54$.

For noncoherent pulses, where every received pulse is independent of each other, we cannot combine the pulses together to form a CW wave as in the coherent case. We have to treat each pulse separately and, therefore, the best we can do in order to improve the accuracy of the frequency estimation is to have an average estimate of all the received pulses. That is,

$$\tilde{f} = \frac{1}{P}\sum_{i=1}^{P} \tilde{f}_i$$

where $\tilde{f}_i$ is the frequency estimate of the $i$th pulse. The variance of the average estimate $\tilde{f}$, is then reduced by a factor of $P$, the number of pulses, i.e.,

$$\mathrm{Var}(\tilde{f}) = \frac{1}{P}\,\frac{6}{(SNR)T_s^2 N(N^2-1)} \tag{70}$$

which implies that the RMS error of the frequency estimate is decreased by $P^{1/2}$. Now, in order to have an accuracy of $10\,Hz$ or less, with each pulse having the values of $\tau = 1\,\mu s$, $SNR = 20\,dB$, and $f_s = 10^8\,Hz$, we need $P \geq$ 151,997 pulses which is too large for practical purposes. Note that an accuracy of the order of $10\,Hz$ is chosen because that is basically what is required in practice for doppler detection and measurement.

One way of reducing the number of pulses $P$ for the same accuracy is to have more samples per pulse, i.e., increase the value of $N$ by sampling at a faster rate, as was explained before. If the number of samples is augmented from $N$ to $QN$, where $Q > 1$, then the MSE is approximately reduced by $Q$ and the RMS error by $Q^{1/2}$. As for the number of noncoherent pulses needed, $P$ is decreased to $P/Q$ and that of coherent pulses to $P/Q^{1/3}$ for the same MSE or RMS error value. In the example of the noncoherent pulses,

if the sampling rate is increased from $10^8$ to $10^9\,Hz$ for the same value of $\tau$, $(Q = 10)$, then for the same accuracy of $10\,Hz$ or less, at least 15,199 pulses are required, which is still very large. Comparing the noncoherent pulses to the case of the coherent ones, where the number of pulses required is in the order of hundred thousand instead of only a few tens, it can be deduced that it is only practically feasible for an ESM system to measure the change in doppler frequency from the coherent pulses and not from the noncoherent pulses, since in general, the number of pulses received during a burst is usually a few hundreds or less.

In the case where the demodulated pulse signal's frequency, $f$ is less than the reciprocal of the pulse duration $\tau$ and is less than the pulse repetition frequency (PRF), i.e., $f < 1/\tau$ and $f < PRF$, each pulse can be taken as a discrete sample with the PRF as the sampling rate. However, this can be done only if the pulse signal is coherent. For a PRF of $10\,KHz$ and a SNR of $20/, dB$, using the expression of the C-R bound in (20), an accuracy of $10\,Hz$ or less can be achieved with at least 12 pulses. For the frequency $f < 1/\tau$ but greater than the PRF ($f > PRF$), each pulse cannot be sampled only once, because this will lead to ambiguities in determining the true frequency. More specifically, there will be a difference of $n(PRF)$ between the true frequency and the estimated frequency, where $n = 1, 2, 3 \cdots$. The three different frequency ranges of the pulse signal from the output of the demodulator are shown in Figure 10. For $f > 1/\tau$, the signal is readily discerned from the information contained in a single pulse, whereas, for $f < 1/\tau$, the pulses are modulated with the amplitude of the signal waveform.

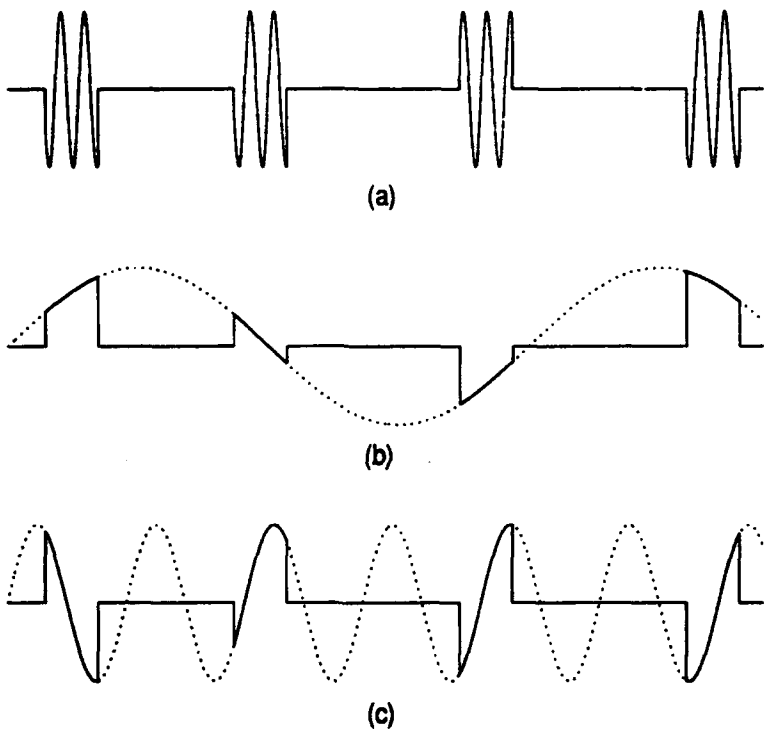


**Figure 10:** (a) pulse train for signal frequency $f > \frac{1}{\tau}$; (b) pulse train for signal frequency $f < \frac{1}{\tau}$ and $f < PRF$; (c) pulse train for signal frequency $f_d < \frac{1}{\tau}$ and $f_d > PRF$.

## REFERENCES


[1] Beyer W.H., *CRC Standard Mathematical Tables and Formulae*, 29th Edition, CRC Press Inc., Boca Raton, Florida, 1991.

[2] Conte S.D. and Carl de Boor, *Elementary Numerical Analysis*, Second Edition, McGraw-Hill Book Company, New York, 1972.

[3] Haykin S., *Adaptive Filter Theory*, Prentice-Hall, Englewood Cliffs, New Jersey, 1986.

[4] Haykin S., *Communication Systems*, John Wiley and Sons Inc., New York, 1983.

[5] Oppenheim A.V. and Schafer R.W., *Discrete-Time Signal Processing*, Prentice-Hall Inc., Englewood Cliffs, New Jersey, 1989.

[6] Palmer L.C., "Coarse Frequency Estimation Using the Discrete Fourier Transform", *IEEE Transaction on Information Theory*, vol. IT-20, p. 104-109, January 1974.

[7] Papoulis A., *Probability Random Variables and Stochastic Processes*, McGraw-Hill Book Company, New York, 1984.

[8] Rife D.C. and Boorstyn R.R., "Single-Tone Parameter Estimation from Discrete-Time Observation", *IEEE Transactions on Information Theory*, vol. IT-20, p. 591-598, September 1974.

[9] Stimson G.W., *Introduction to Airborne Radar*, Hughes Aircraft Company, El Segundo, California, 1983.

[10] Van Trees H. L., *Detection, Estimation and Modulation Theory*, Part I, John Wiley and Sons Inc., New York 1968.

[11] Whalen A.D., *Detection of Signals in Noise*, Academic Press, New York, 1971.

[12] Wiley R.G., *Electronic Intelligence: The Analysis of Radar Signals*, Artech House Inc., Norwood, MA., 1982.

## DOCUMENT CONTROL DATA

(Security classification of title, body of abstract and indexing annotation must be entered when the overall document is classified)

1. ORIGINATOR (the name and address of the organization preparing the document. Organizations for whom the document was prepared, e.g. Establishment sponsoring a contractor's report, or tasking agency, are entered in section 8.)

DEFENCE RESEARCH ESTABLISHMENT OTTAWA
NATIONAL DEFENCE
SHIRLEY BAY, OTTAWA, ONTARIO K1A 0K2 CANADA

2. SECURITY CLASSIFICATION (overall security classification of the document. including special warning terms if applicable)

UNCLASSIFIED

3. TITLE (the complete document title as indicated on the title page. Its classification should be indicated by the appropriate abbreviation (S,C or U) in parentheses after the title.)

ESTIMATION OF DOPPLER FREQUENCY IN THE PRESENCE OF NOISE (U)

4. AUTHORS (Last name, first name, middle initial)

LIE CHIN CHEONG, PATRICK

5. DATE OF PUBLICATION (month and year of publication of document)

SEPTEMBER 1993

6a. NO. OF PAGES (total containing information. Include Annexes, Appendices, etc.)

40

6b. NO. OF REFS (total cited in document)

12

7. DESCRIPTIVE NOTES (the category of the document, e.g. technical report, technical note or memorandum. If appropriate, enter the type of report, e.g. interim, progress, summary, annual or final. Give the inclusive dates when a specific reporting period is covered.)

DREO TECHNICAL NOTE

8. SPONSORING ACTIVITY (the name of the department project office or laboratory sponsoring the research and development. Include the address.)

DEFENCE RESEARCH ESTABLISHMENT OTTAWA
NATIONAL DEFENCE
SHIRLEY BAY, OTTAWA, ONTARIO K1A 0K2 CANADA

9a. PROJECT OR GRANT NO. (if appropriate, the applicable research and development project or grant number under which the document was written. Please specify whether project or grant)

011LB11

9b. CONTRACT NO. (if appropriate, the applicable number under which the document was written)

10a. ORIGINATOR'S DOCUMENT NUMBER (the official document number by which the document is identified by the originating activity. This number must be unique to this document.)

DREO TECHNICAL NOTE 93-15

10b. OTHER DOCUMENT NOS. (Any other numbers which may be assigned this document either by the originator or by the sponsor)

11. DOCUMENT AVAILABILITY (any limitations on further dissemination of the document, other than those imposed by security classification)

(XX) Unlimited distribution
( ) Distribution limited to defence departments and defence contractors; further distribution only as approved
( ) Distribution limited to defence departments and Canadian defence contractors; further distribution only as approved
( ) Distribution limited to government departments and agencies; further distribution only as approved
( ) Distribution limited to defence departments; further distribution only as approved
( ) Other (please specify):

12. DOCUMENT ANNOUNCEMENT (any limitation to the bibliographic announcement of this document. This will normally correspond to the Document Availabilty (11). However, where further distribution (beyond the audience specified in 11) is possible, a wider announcement audience may be selected.)

13. ABSTRACT ( a brief and factual summary of the document. It may also appear elsewhere in the body of the document itself. It is highly desirable that the abstract of classified documents be unclassified. Each paragraph of the abstract shall begin with an indication of the security classification of the information in the paragraph (unless the document itself is unclassified) represented as (S), (C), or (U). It is not necessary to include here abstracts in both offical languages unless the text is bilingual).

(U) This paper discusses the estimation of the Doppler frequency of an incoming RF signal that is corrupted by additive white Gaussian noise. The estimation is performed using the sampled in-phase and quadrature components of the demodulated signal. The Cramér-Rao lower bounds for the minimum variance estimators as well as the maximum-likelihood (ML) estimator are derived. The mean-square error of the ML estimator is then studied.

14. KEYWORDS, DESCRIPTORS or IDENTIFIERS (technically meaningful terms or short phrases that characterize a document and could be helpful in cataloguing the document. They should be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location may also be included. If possible keywords should be selected from a published thesaurus. e.g. Thesaurus of Engineering and Scientific Terms (TEST) and that thesaurus-identified. If it is not possible to select indexing terms which are Unclassified, the classification of each should be indicated as with the title.)

DOPPLER FREQUENCY ESTIMATION
CRAMER-RAO BOUNDS FOR MINIMUM VARIANCE UNBIASED ESTIMATOR
MAXIMUM LIKELIHOOD
FAST FOURIER TRANSFORM
POLYNOMIAL PHASE ESTIMATES